Negative effects of psychotherapies for adult depression

Cuijpers, Pim; Reijnders, Mirjam; Karyotaki, Eirini; de Wit, Leonore; Ebert, David D.

published in
Journal of Affective Disorders
2018

DOI (link to publisher)
10.1016/j.jad.2018.05.050

document version
Publisher's PDF, also known as Version of record

citation for published version (APA)
Cuijpers, P., Reijnders, M., Karyotaki, E., de Wit, L., & Ebert, D. D. (2018). Negative effects of psychotherapies for adult depression: A meta-analysis of deterioration rates. *Journal of Affective Disorders*, *239*, 138-145. https://doi.org/10.1016/j.jad.2018.05.050

Journal of Affective Disorders 239 (2018) 138–145

Contents lists available at ScienceDirect

Journal of Affective Disorders

journal homepage: www.elsevier.com/locate/jad

Review article

Negative effects of psychotherapies for adult depression: A meta-analysis of deterioration rates

Pim Cuijpers[a,*], Mirjam Reijnders[a], Eirini Karyotaki[a], Leonore de Wit[a], David D. Ebert[b]

[a] Department of Clinical, Neuro and Developmental Psychology, Amsterdam Public Health research institute, Vrije Universiteit Amsterdam, The Netherlands
[b] Clinical Psychology and Psychotherapy, Institute for Psychology, Friedrich-Alexander University Erlangen-Nuremberg, Erlangen, Germany

ARTICLE INFO

Keywords:
Depression
Negative outcomes
Deterioration rates
Psychotherapy
Cognitive behavior therapy
Meta-analysis

ABSTRACT

Background: The risk for deterioration in patients receiving psychotherapy for adult depression has not been examined extensively and it is not clear whether psychotherapy reduces this risk or may even increase it in some patients. We conducted a meta-analysis of trials comparing these psychotherapies with control conditions that report deterioration rates.

Methods: We used an existing database of randomized trials on psychotherapies for adult depression which was updated up to 1/1/2017, through systematic searches in bibliographic databases. We included trials that reported clinically significant deterioration rates.

Results: We included 18 studies with 23 comparisons between therapy and control groups. The pooled risk ratio of deterioration was 0.39 (95% CI: 0.27 ~ 0.57), indicating that patients in the psychotherapy groups have a 61% lower chance to deteriorate than patients in the control groups. We found that 20 patients need to be treated with psychotherapy in order to avoid one case of deterioration, compared to the control conditions. The median deterioration rate in the therapy groups was 4%, and in some studies more than 10%, indicating that clinicians should always be aware of the risk of deterioration.

Limitations: The results should be considered with caution because most studies had at least some risk of bias. Only 6% of all trials comparing psychotherapy with a control condition reported deterioration rates, using different ways to define deterioration which made pooling the prevalence rates across treatments and control groups impossible.

Conclusions: Psychological treatments of adult depression may reduce the risk for deterioration, compared to control groups, but this should be considered with caution because of the small proportion of studies reporting deterioration rates.

1. Introduction

It is well-established that several types of psychotherapies are effective in the treatment of adult depression, including cognitive behavior therapy (Cuijpers et al., 2016a), behavioral activation therapy (Ekers et al., 2014), interpersonal psychotherapy (Cuijpers et al., 2016b), problem-solving therapy (Malouff et al., 2007), and possibly brief psychodynamic therapy (Driessen et al., 2015) and non-directive counseling (Cuijpers et al., 2012). There is also considerable evidence that the effects of these therapies do not or only marginally differ from each other (Barth et al., 2013).

Apart from the average positive effects these therapies can have on depressed patients, it is also important to examine whether these therapies may have negative effects on some individual patients (Barlow, 2010; Foulkes, 2010; Peterson et al., 2013). Although the importance of negative effects of psychotherapies has been described for several decades (Hadley and Strupp, 1976; Mohr, 1995), it is relatively recent that this is considered one of the core issues that should be prioritized in research on psychotherapies (Barlow, 2010; Lilienfeld, 2007). At the moment, it can be said that there is consensus in the field of psychotherapy research that (1) negative effects should be examined better and (2) has mostly been neglected in much of this research up to now (Barlow, 2010; Dimidjian and Hollon, 2010).

How negative effects should be defined is less clear (Boisvert, 2010; Dimidjian and Hollon, 2010). It is clear that an increased risk of deterioration during therapy is one of the core types of negative effects. However, there are other types of negative effects that are also important to consider, such as serious adverse events (Rozental et al.,

* Corresponding author.
E-mail address: p.cuijpers@vu.nl (P. Cuijpers).

https://doi.org/10.1016/j.jad.2018.05.050
Received 11 August 2017; Received in revised form 10 April 2018; Accepted 28 May 2018
Available online 03 July 2018

2014). Also non-response and drop-out can be considered as negative effects as they could have prevented the patient from receiving adequate care or spontaneous remission (Dimidjian and Hollon, 2010). In this paper we will focus on the negative effects of psychotherapies in terms of clinically relevant deterioration, because this is one of the most important type of negative outcomes, it has not been examined in conventional meta-analyses before.

Although there is a considerable number of randomized trials that report deterioration rates, to the best of our knowledge these have not been integrated in meta-analytic research. There are some "individual patient data" meta-analyses that have reported deterioration rates in psychotherapies (Ebert et al., 2016; Vittengl et al., 2016). However, these meta-analyses are aimed at only one type of treatment (internet-based treatment of depression; (Ebert et al., 2016)), or focus on the difference in deterioration rates between cognitive behavior therapy and pharmacotherapy for depression (Vittengl et al., 2016). No meta-analysis has examined deterioration rates in psychotherapy versus untreated control groups.

One important reason why no meta-analysis on deterioration rates in psychotherapies has been conducted is that these deterioration rates are typically not reported in titles and abstracts of studies, but only reported in the full papers because this is typically not the primary outcome of trials (only one of the 18 studies included in the current meta-analysis reported negative outcomes in the abstract; (Hautzinger et al. 2004)). At the same time only a relatively small proportion of trials do report deterioration rates. Therefore, all papers on trials in a certain field have to be collected and examined to determine whether they report relevant data. In the current study we solved this by using an existing database of randomized trials on psychotherapies for adult depression that is updated every year and record whether they report deterioration rates.

It has been estimated that 5–10% of patients deteriorate during therapy (Lambert, 2007). However, that does not necessarily have to be the result of the therapy, but may be related to other causes (Dimidjian and Hollon, 2010). Therefore when assessing deterioration rates in psychotherapy, it is important to compare these rates with those in control groups.

In the current meta-analysis, we selected studies in which psychotherapies for adult depression were compared with control conditions (waiting list, care as usual, placebo, other inactive control group), and examined whether they reported the number of patients who deteriorated (using any measure for deterioration). We then pooled the results of these studies to estimate whether psychotherapy resulted in lower or higher deterioration rates than the control conditions. We also compared the characteristics of the patients, therapies and designs of the studies that report deterioration rates with those that do not.

2. Methods

2.1. Identification and selection of studies

We used an existing database of studies on the psychological treatment of depression. This database has been described in detail elsewhere (Cuijpers et al., 2008b), and has been used in a series of earlier published meta-analyses (Cuijpers, 2017). For this database we searched four major bibliographical databases (PubMed, PsycInfo, Embase and the Cochrane Library) by combining terms (both index terms and text words) indicative of depression and psychotherapies, with filters for randomized controlled trials. The full search string for one database (PubMed) is given in Appendix A. We also searched a number of bibliographical databases to identify trials in non-Western countries (Cuijpers et al., 2018), because the number of trials on psychological treatments in these countries is growing rapidly. Furthermore, we checked the references of earlier meta-analyses on psychological treatments of depression. The database is continuously updated and was developed through a comprehensive literature search (from 1966 to January 1st 2017). All records were screened by two independent researchers and all papers that could possibly meet inclusion criteria according to one of the researchers were retrieved as full-text. The decision to include or exclude a study in the database was also done by the two independent researchers, and disagreements were solved through discussion.

We included studies that were: (a) a randomized trial (b) on a psychological treatment (c) for adult depression that was (d) compared with a control group (waiting list, care-as-usual, placebo, other inactive treatment) (e) and reported deterioration rates for the psychological treatment and the control group. Depression could be established with a diagnostic interview or with a score above a cut-off on a self-report measure. Psychotherapy was defined as an intervention with a primary focus on language-based communication between a patient and a therapist, or as bibliotherapy supported by a therapist (Barth et al., 2013). We allowed any definition of clinically significant deterioration, as long as it indicated the proportion of patients in therapy and control groups who scored higher on depression symptom severity after treatment than they did at baseline, and as long as the authors described this as an indication for clinically significant deterioration. Co-morbid mental or somatic disorders were not used as an exclusion criterion. Studies on inpatients were excluded. We also excluded maintenance studies, aimed at people who had already recovered or partly recovered after an earlier treatment. In order to compare the characteristics of the studies that reported deterioration rates and those that did not, we also included studies that met all inclusion criteria, except (e) and did not report deterioration rates.

2.2. Quality assessment and data extraction

We assessed the validity of included studies using four criteria inspired by the 'Risk of bias' assessment tool, developed by the Cochrane Collaboration (Higgins et al., 2011). This tool assesses possible sources of bias in randomized trials, including the adequate generation of allocation sequence; the concealment of allocation to conditions; the prevention of knowledge of the allocated intervention (masking of assessors); and dealing with incomplete outcome data (this was assessed as positive when intention-to-treat analyses were conducted, meaning that all randomized patients were included in the analyses). Assessment of the validity of the included studies was conducted by two independent researchers, and disagreements were solved through discussion.

We also coded participant characteristics (recruitment: community, clinical, other; diagnosis: according to a diagnostic interview or a self-report measure; target group: adults in general or a specific target group, such as older adults, patient with a comorbid general medical disorder or women with postpartum depression); characteristics of the psychotherapies (type of therapy: CBT versus another therapy; treatment format: individual, group, guided self-help, other; number of sessions); and general characteristics of the studies (type of control group: waiting list, care as usual, other; risk of bias; year of publication).

2.3. Outcome measures

For each comparison between a psychotherapy and a control condition, we calculated the risk ratio (RR), indicating the proportion of patients that deteriorated in the therapy group, divided by the proportion in the control group. A RR of 1 indicates that the proportion of deteriorated patients in the therapy group is the same as the proportion in the control group, a RR below 1 indicates that the proportion of deteriorated patients is lower in the treatment group. Conversely, a RR above 1 indicates the proportionis higher in the treatment group.

We also calculated the risk difference (RD), which indicates the difference between the proportion of deterioration in the therapy group and the control group. We used the RD to calculate the numbers-

needed-to-be-treated (NNT), which indicates the number of patients that have to be treated in order to avoid one case of deterioration (the NNT is 1 divided by the RD).

2.4. Meta-analyses

To calculate pooled RRs and RDs, we used the computer program Comprehensive Meta-Analysis (version 3.3070; CMA). Because we expected considerable heterogeneity among the studies, we employed a random effects pooling model in all analyses. As a test of homogeneity of effect sizes, we calculated the I^2-statistic, which is an indicator of heterogeneity in percentages. A value of 0% indicates no observed heterogeneity, and larger values indicate increasing heterogeneity, with 25% as low, 50% as moderate, and 75% as high heterogeneity (Higgins et al., 2003). We calculated 95% confidence intervals around I^2 (Ioannidis et al., 2007), using the non-central chi-squared-based approach within the heterogi module for Stata (Orsini et al., 2006).

We conducted subgroup analyses according to the mixed effects model, in which studies within subgroups are pooled with the random effects model, while tests for significant differences between subgroups are conducted with the fixed effects model. Because the number of studies was relatively small, we conducted only subgroup analyses for the most basic characteristics of the patients (whether or not depression was established with a diagnostic interview), the interventions (type of therapy: CBT versus other therapies; format: individual, group or guided self-help) and the studies (type of control group: waiting list, care as usual, other), as well as the definition of deterioration..

We conducted sensitivity analyses to examine whether risk of bias was associated with the outcomes. We compared studies with low risk of bias (no bias on any of the four items) to studies with risk of bias according to at least one of the items.

For continuous variables, we used meta-regression analyses to test whether there was a significant relationship between the continuous variable and effect size, as indicated by a Z-value and an associated p-value. Multivariate meta-regression analyses, with the effect size as the dependent variable, were conducted in CMA.

We tested for publication bias by inspecting the funnel plot on primary outcome measures and by Duval and Tweedie's trim and fill procedure (Duval and Tweedie, 2000), which yields an estimate of the effect size after the publication bias has been taken into account (as implemented in CMA). We also conducted Egger's test of the intercept to quantify the bias captured by the funnel plot and to test whether it was significant.

2.5. Comparisons between studies that did and did not report deterioration rates

In order to examine whether the participants, interventions and general characteristics of the studies reporting deterioration rates differed from that did not report them, we compared basic characteristics of the two samples. We compared the two samples of studies with chi-square tests for categorical variables and with *t*-tests for continuous variables. These analyses were done with SPSS, version 23.

3. Results

3.1. Selection and inclusion of studies

After examining a total of 18,500 abstracts during the building of the database described in the Methods section (14,290 after removal of duplicates), we retrieved 2092 full-text papers for further consideration. We excluded 2074 of the retrieved papers. The PRISMA flowchart describing the inclusion process, including the reasons for exclusion, is presented in Fig. 1. A total of 18 studies with 23 comparisons between a therapy and a control group met inclusion criteria for this meta-analysis (Table 1).

For the comparison between studies that reported deterioration rates and those that did not, we included another 271 studies with 346 comparisons between a therapy and a control group, that met all inclusion criteria, except that deterioration rates were not included. The references and characteristics of the studies are given in another paper (Cuijpers et al., 2018). Among the total sample of studies ($N = 289$) only 6.2% ($N = 18$) reported deterioration rates.

3.2. Characteristics of included studies

Selected characteristics of the included studies are presented in Table 1. The 18 studies with 23 comparisons between a therapy and a control group included a total of 1655 patients (519 in the CBT intervention groups, 350 in the other intervention groups, 329 in the waiting list control groups, 294 in the care-as-usual control groups and 163 in the other control groups). A total of six studies were aimed at adults in general and 12 at other more specific target groups. Twelve studies recruited patients (also) from the community, three recruited exclusively from clinical populations, three used other recruitment methods. Nine studies with 12 comparisons used a diagnostic measure to diagnose a depressive disorder, while nine studies with 11 comparisons used clinical cut-off scores. Treatment was delivered in individual format in five studies with six comparisons, in group format in seven studies with 11 comparisons, and in guided self-help format in six studies with six comparisons. The number of treatment sessions ranged from six to 12. In 10 studies with 13 comparisons a waiting list control group was used, five studies with five comparisons had a care-as-usual control group, and three studies with five comparisons used another control group. Five studies were conducted in the United States of America, two in Canada, three in Australia, three in Germany, two in Sweden, one in China, one in Finland and one in the United Kingdom. Deterioration was measured using the criteria for clinically significant deterioration by Jacobson and Truax (1991) in 11 studies, while 7 studies used another definition (such as moving to another symptom category or scoring a number of points lower on a depression scale).

3.3. Quality ratings

The quality of the studies varied. Nine studies reported an adequate sequence generation, while the other nine did not. Six studies reported allocation to conditions by an independent (third) party. Five studies reported blinding of outcome assessors while another 12 used only self-report outcomes. A total of 11 studies conducted intention-to-treat analyses. Only three studies met all four quality criteria, seven met three criteria, two met two criteria, and the remaining six studies met one of the criteria.

3.4. Deterioration rates of psychotherapies compared to control groups

The pooled risk ratio (RR) of deterioration rates in all 23 comparisons between a psychotherapy and control group was RR = 0.39 (95% CI: 0.27–0.57), with zero heterogeneity ($I^2 = 0$; 95% CI: 0–40). This means that the risk for deterioration in the psychotherapy groups was reduced by 61% compared to the control groups. The risk difference (RD) between psychotherapy and control groups, which indicates the risk of deterioration in the control groups minus the risk in the psychotherapy groups, was RD = 0.05 (95% CI: 0.03–0.07), which corresponds with a numbers-needed-to-be-treated (NNT) of 20. The forest plot with the RRs, RDs and the exact number of deteriorated patients in each study is presented in Fig. 2. The deterioration rates in the psychotherapy groups ranged from 0 to 25%, with a median of 4%. The deterioration rates in the control conditions ranged from 1 to 44% with a median of 11%.

As can be seen from Fig. 2, the majority of trials did not result in significant differences between the treatment and control groups in terms of deterioration. Only three of the 23 comparisons were

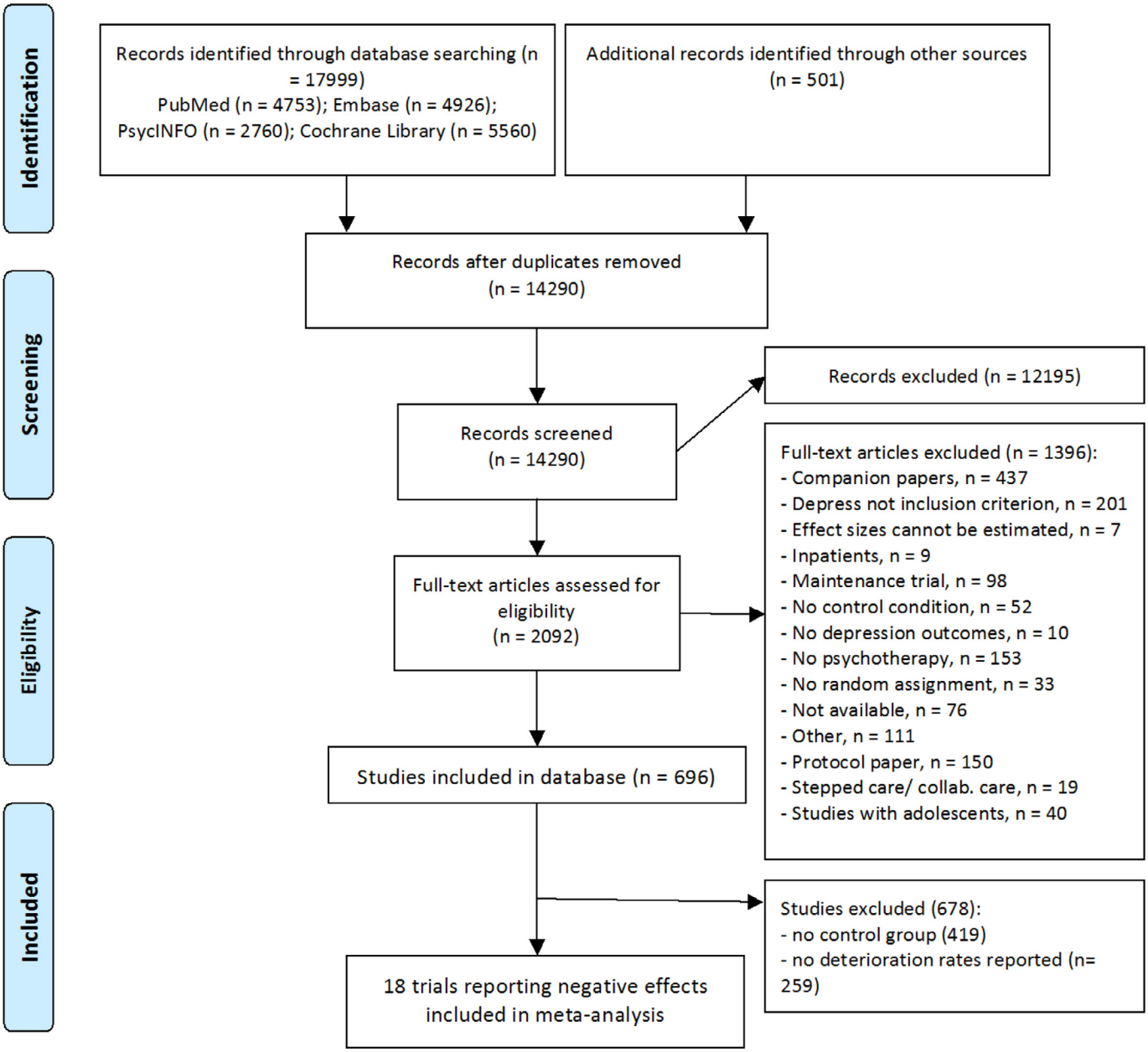


Fig. 1. Flowchart for the inclusion of studies.

significant, which is not surprising considering the small number of patients that deteriorated and low statistical power in each study.

In four studies, two or more types of therapies were compared with the same control group (three studies with two comparisons, and one with three comparisons). Because these effect sizes were not independent of each other, they may have artificially reduced heterogeneity and influenced the outcomes. We conducted two separate analyses to examine this. In the first analysis we only included the largest effect size from each study in the pooling of the outcomes, and in the second only the smallest effect size. As can be seen in Table 2, the effect sizes and levels of heterogeneity were comparable to those in the main analyses.

We found no indication for publication bias. Duvall and Tweedie's trim and fill procedure did not indicate that studies were missing (the adjusted RR was identical to the original RR), and Egger's test of the intercept was not significant ($p > 0.1$).

We conducted a series of subgroup analyses to examine possible associations of the RR with characteristics of the participants, therapies and the studies. Because the number of studies was not large, we limited these analyses to some of the most essential ones (Table 2). We found no indications that the RR was associated with the definition of depression (according to a diagnostic interview versus a high score on a self-rating depression measure), type of therapy (CBT versus others), treatment format (individual, group, guided self-help), and type of control group (waiting list, care as usual, other).

We examined risk of bias in sensitivity analyses in which we examined potential differences between studies with no risk of bias compared to studies with any risk of bias (Table 2). We found no indication that studies with low risk of bias had different outcomes than the other studies, although the number of studies with low risk of bias was small and not finding a significant difference may be related to low statistical power.

Table 1
Selected characteristics of randomized trials of psychotherapies for adult depression reporting deterioration rates ($N = 18$).

Study	Type	Recr	Target grp	Diagn	Conditions	N	Nsess	Format	Country	RoB[a]
Buntrock et al. (2015)	CBT	Comm	Adults	Subthr. Depr.; CES-D ≥ 16	iCBT CAU	202 204	6	gsh	Germany	+ + + +
Carlbring et al. (2013)	Other	Comm	Adults	MDD: DSM-IV, SCID	iBA + ACT WL	40 40	7	gsh	Sweden	+ + SR +
Choy (2016)	Other	Other	Elderly	Cut–off; GDS–15–Ch 8–13	Instr. RT WL	39 33	8	grp	China	– – SR –
Gitlin et al. (2013)	Other	Comm	Elderly	Cut-off; PHQ-9 ≥ 5	Home-B I WL	89 93	10	ind	USA	+ + SR +
Hallford (2016)	Other	Clin	Adults	Cut-off; DASS-21-SF ≥ 7	CRT CAU	14 12	6	ind	Australia	– + SR +
Hautzinger (2004)	CBT	Comm	Elderly	Unipolar Depr; DSM-IV, SCID	CBT WL	65 35	12	grp	Germany	+ – SR +
Kelly et al. (1993)	CBT/SUP	Comm	Gen. Med.	Cut-off; CES-D ≥ 16	CBT SUP Comp. Cond.	27 14 27	8	grp	USA	– – SR –
Lappalainen et al. (2015)	Other	Comm	Adults	MDE; DSM-IV-TR	iACT WL	18 20	6	gsh	Finland	– + SR +
Mulcahy et al. (2010)	IPT	Clin	Women with PPD	MDD; DSM-IV, MCMI-III	IPT CAU	23 27	11	grp	Australia	+ – – –
Nobis et al. (2015)	CBT	Comm	Gen. Med.	Cut-off; CES-D ≥ 23	iCBT Psychoed.	129 127	6	gsh	Germany	+ – SR +
Pugh et al. (2015)	CBT	Comm	Women with PPD	Cut-off; EPDS ≥ 10	iCBT WL	24 21	12	gsh	Canada	+ + SR –
Ransom et al. (2008)	IPT	Comm	Gen. Med.	MDE / Dysth.; PRIME-MD	IPT CAU	31 35	6	ind	USA	– – SR +
Scott and Stradling (1990)	CBT	Clin	Adults	Primary / Probable MDD; RDC, PSE	CBT – gct CBT – ict WL	10 19 19	12	Grp ind	UK	– – SR +
Segre et al. (2015)	Other	Other	Other	Cut-off; EPDS ≥ 12	LV WL	41 25	6	ind	USA	+ – + +
Watt and Cappeliez (2000)	Other	Comm	Elderly	Cut-off; GDS ≥ 14	Instr. RT Integr. RT Active Soc.	9 9 9	6	grp	Canada	– – + –
Wickberg and Hwang (1996)	SUP	Other	Women with PPD	MDD; DSM-III-R, MADRS	SUP CAU	15 16	6	ind	Sweden	– – + –
Wollersheim and Wilson (1991)	CBT	Comm	Adults	MDD; DSM-III, MMPI ≥ 70	CBT – cop CBT – gsh SUP WL	8 8 8 8	10	Grp gsh grp	USA	– – SR –
Wuthrich and Rapee (2013)	CBT	Comm	Elderly	Anx/Mood Dis; DSM-IV, ADIS	CBT WL	27 35	12	grp	Australia	+ – + +

Abbreviations: ACT: acceptance and commitment therapy; Anx: anxiety; AU: Australia; CAN: Canada; CAU: care as usual; CBT: cognitive behavioral therapy; CES-D: center for epidemiological studies depression scale; Clin: participants recruited in a clinical setting; Comm: participants recruited in a community setting; Comp. Cond.: comparison condition; CRT: cognitive-reminiscence therapy; CT: cognitive therapy; DASS-21-SF: depression anxiety and stress scale 21-item short-form; Diagn: diagnosis; Dis: disorder; Dysth.: dysthymic disorder; EU: Europe; GDS: geriatric depression scale; GDS-15-Ch: geriatric depression scale short form Chinese version; Gen. Med.: general medical disorder; Grp: group format; Gsh: guided self-help format; Home-B I: home-based intervention; iACT: web-based acceptance and commitment-based therapy; iBA: internet-based behavioral activation; iCBT: web-based cognitive behavioral intervention; Ind: individual format; Instr. RT: instrumental reminiscence therapy; Integr. RT: integrative reminiscence therapy; LV: listening visits; MCMI-III: Millon clinical multiaxial inventory-III; MDD: major depressive disorder; MDE: major depressive episode; Nsess: number of sessions; PHQ-9: patient health questionnaire; PPD: post-partum depression; PRIME-MD: primary care evaluation of mental disorders, Mood Module; PSE: present state examination; Psychoed: psychoeducation; Recr: recruitment; RDC: Spitzer's research diagnostic criteria; RoB: risk of bias; Soc: socialization; Subthr. Depr.: subthreshold depression; SUP: nondirective supportive therapy; UK: United Kingdom; USA: United States of America; WL: waitlist.

[a] In this column a positive (+) or negative (−) sign is given for four quality criteria of the study, respectively: allocation sequence; concealment of allocation to conditions; blinding of assessors; and intention-to-treat analyses. Sr in the third criterion indicates that only self-report measures (and no assessor) were used.

3.5. Differences between studies that did and did not report deterioration rates

The characteristics of the participants, therapies and design of the studies that reported deterioration rates and those that did not are reported in Table 3. There were no significant differences between the two samples of studies.

4. Discussion

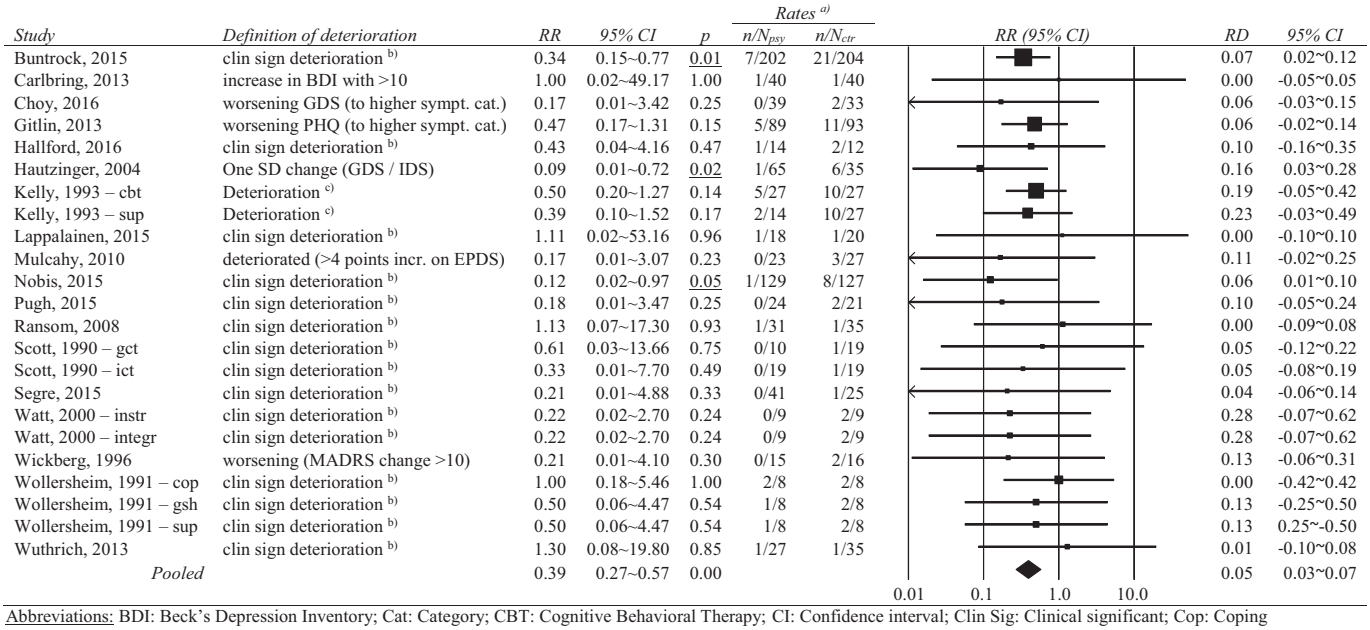
In this meta-analysis we pooled the deterioration rates reported in randomized trials of psychotherapies for adult depression. We found that psychotherapy significantly reduces deterioration rates of patients compared to patients in control conditions. Overall, we found that the risk of deterioration in the psychotherapy groups was reduced by 61% compared to untreated control conditions. We found that 20 patients need to be treated with psychotherapy in order to avoid one case of deterioration, compared to the control conditions.

Heterogeneity was zero in the studies reporting deterioration rates, suggesting that all studies point in the same direction, although the 95% CI around the level of heterogeneity is too large to draw definite conclusions about that. The low level of heterogeneity is in contrast with other meta-analyses examining the effects of psychotherapies on depressive symptoms (e.g., Cuijpers et al., 2013; Cuijpers et al., 2016a). However, the number of trials reporting deterioration rates was low and it is not clear whether the studies that do report these rates are representative of all trials on psychotherapy for depression. Overall, only 6% of all trials comparing psychotherapy for depression with a control group, reported the deterioration rate.

That only 6% of trials report deterioration rates is concerning. The need for research on negative effects has been acknowledged for several years now (Barlow, 2010; Foulkes, 2010; Peterson et al., 2013), but even in recent trials negative outcomes are usually not reported. In the 141 randomised controlled trials on psychotherapy for adult depression, only 8% reported deterioration rates.

One reason why authors do not report negative outcomes in trials may be that these outcomes do not significantly differ between treatment and control groups. Because usually only a small numbers of patient deteriorate, there is typically not enough statistical power to detect significant differences between treatment and control groups. Authors may think that it is not important enough to report such outcomes because of the non-significance. However, whether or not outcomes are significant should not be leading in reporting outcomes of trials. Even if differences between therapy and control groups on negative outcomes are not significant, they should be reported in randomized trials, because these outcomes can be pooled in meta-analyses and still result in important knowledge about the negative outcomes and predictors of such outcomes.

While the mean risk for deterioration was reduced compared to the control group, it may nevertheless be the case that patients with certain



Abbreviations: BDI: Beck's Depression Inventory; Cat: Category; CBT: Cognitive Behavioral Therapy; CI: Confidence interval; Clin Sig: Clinical significant; Cop: Coping Therapy; EPDS: Edinburgh Postnatal Depression Scale; Gct: Group Cognitive Therapy; GDS: Geriatric Depression Scale; Gsh: Guided self-help; HAMD: Hamilton Rating Scale for Depression; Ict: Individual Cognitive Therapy; Instr: Instrumental; Integr: Integrative; MADRS: Montgomery-Asberg Depression Rating Scale; RD: risk difference; RR: risk ratio; Sup: Supportive Therapy.

a)These rates indicate the number of patients who deteriorated (n) and the total psychotherapy and control group (N)

b) Clinical significant deterioration according to the definition of Jacobson & Truax (1991)

c) According to Speer & Swondle (1982).

Fig. 2. Forest plot of deterioration rates in studies comparing psychotherapy for adult depression with control groups: Risk ratio's.

characteristics experience a deterioration as a consequence of the therapy. Individual patient data (IPD) meta analyses can conduct analyses that have not been reported by the primary studies. For example, one recent IPD meta-analysis of trials on internet-based treatments of depression also found lower deterioration rates for patients receiving treatment compared to control groups, but also found that patients with lower education had a higher risk for deterioration than patients with higher education (Ebert et al., 2016). Hence, future studies should use such approaches in order to identify subgroups that might experience a worse adjustment as a result of the treatment.

The results of this study are in line with a growing number of meta-analytic studies showing that psychological treatments result in a significant reduction of deterioration in patients. Earlier studies were limited to small samples (Vittengl et al., 2016), and one type of intervention (internet-based CBT; (Ebert et al., 2016)). However, these studies also support the findings that deterioration rates are low and that psychological treatment reduces these rates.

Deterioration rates are low in most studies, but it depends on how it is measured. In some studies the deterioration rates are higher than 10% in the therapy groups. That means that deterioration is something that therapists should always be cautious about. It cannot be expected that therapies work for everyone and that all patients improve during therapy. Therapists should be trained in recognition and evaluation of deterioration and other negative effects of therapy, as well as how to

Table 2
Deterioration rates of psychotherapies for adult depression compared with control groups: risk ratios, risk differences and numbers-needed-to-be-treated[a].

		Nc	RR	95% CI	I^2	95% CI	p	RD	95% CI	NNT
All comparisons		23	0.39	0.27–0.57	0	0–40		0.05	0.03–0.07	20
One comparison per study (only highest)		18	0.35	0.22–0.54	0	0–44		0.05	0.03–0.07	20
One comparison per study (only lowest)		18	0.38	0.25–0.58	0	0–44		0.05	0.03–0.07	20
Subgroup analyses										
Definition of deterioration	Jacobson and Truax (1991)	15	0.39	0.23–0.66	0	0–46	0.94	0.05	0.02–0.07	20
	Other	8	0.38	0.22–0.66	0	0–56		0.08	0.02–0.13	13
Diagnosis	Depressive disorder	12	0.48	0.23–1.01	0	0–50	0.51	0.02	−0.01–0.05	50
	Above cut-off	11	0.36	0.23–0.56	0	0–51		0.07	0.04–0.09	14
Type of therapy	CBT	11	0.39	0.24–0.63	0	0–51	0.98	0.06	0.03–0.09	17
	Other	12	0.39	0.21–0.71	0	0–50		0.04	0.01–0.07	25
Format	Individual	6	0.44	0.20–0.97	0	0–61	0.82	0.04	0.00–0.09	25
	Group	11	0.41	0.24–0.71	0	0–51		0.08	0.03–0.13	13
	Guided self-help	6	0.33	0.16–0.65	0	0–61		0.04	0.01–0.07	25
Control group	Waiting list	13	0.44	0.24–0.81	0	0–49	0.85	0.03	0.00–0.06	33
	Care as usual	5	0.35	0.17–0.71	0	0–64		0.06	0.02 to −0.10	17
	Other	5	0.36	0.19–0.71	0	0–64		0.13	0.03–0.23	8
Sensitivity analyses: risk of bias										
Risk of bias	Low	6	0.36	0.20–0.64	0	0–61	0.70	0.05	0.01–0.09	20
	At least some RoB	17	0.41	0.25–0.68	0	0–45		0.05	0.02–0.08	20

Abbreviations: CBT: cognitive behavioral therapy; CI: confidence interval; N_c: number of comparisons; RD: risk difference; RR: risk ratio; NNT: numbers-needed-to-treat; RoB: risk of bias.

[a] according to the random effects model.

Table 3
Differences between randomised controlled trials on psychotherapy for adult depression that report negative effects (N = 23) and those that do not (*N* = 346) report negatieve effects.

		Studies with negative effects	Studies without negative effects	All studies	p-value[a)]
		% (N)	% (N)	% (N)	
Recruitment	Community	45.1 (156)	69.6 (16)	46.6 (172)	0.06
	Clinical	22.8 (79)	17.4 (4)	22.5 (83)	
	Other	32.1 (111)	13.0 (3)	30.9 (114)	
Diagnosis	Diagnostic interview	52.9 (183)	52.2 (12)	52.8 (195)	0.95
	Self-report measure	47.1 (163)	47.8 (11)	47.2 (174)	
Target group	Adults	53.5 (185)	60.9 (14)	53.9 (199)	0.49
	Specific group	46.5 (161)	39.1 (9)	46.1 (170)	
Therapy	CBT	54.9 (190)	34.8 (8)	53.7 (198)	0.06
	Other	45.1 (156)	65.2 (15)	46.3 (171)	
Format	Individual	45.1 (156)	26.1 (6)	43.9 (162)	0.14
	Group	32.9 (114)	47.8 (11)	33.9 (125)	
	Guided self-help	16.8 (58)	26.1 (6)	17.3 (64)	
	Other	18 (5.2)	0.0 (0)	4.9 (18)	
Number of sessions	≤8	54.2 (186)	56.5 (13)	54.4 (199)	0.06
	9–12	28.3 (97)	43.5 (10)	29.2 (107)	
	≥13	17.5 (60)	0.0 (0)	16.4 (60)	
Control group	Waiting list	42.2 (146)	56.5 (13)	43.1 (159)	0.21
	Care as usual	40.2 (139)	21.7 (5)	39.0 (144)	
	Other	17.6 (61)	21.7 (5)	17.9 (66)	
Risk of bias	0–2	54.3 (188)	47.8 (11)	53.9 (199)	0.80
	3–4	45.7 (158)	52.2 (12)	46.1 (170)	
Year of publication	≤1990	17.6 (61)	8.7 (2)	17.1 (63)	
	1991–2000	13.0 (45)	34.8 (8)	14.4 (53)	0.91
	2001–2010	26.6 (92)	13.0 (3)	25.7 (95)	
	≥2011	42.8 (148)	43.5 (10)	42.8 (158)	

[a)] Differences between studies were tested with Chi-square tests for all variables, except for number of sessions, risk of bias and year of publication; these are continuous variables and therefore tested with *t*-tests.

plan treatment taking possible negative consequences into account (Linden and Schermuly-Haupt, 2014).

One of the major problems of this study was that deterioration was measured in many different ways. It is important for future research to develop consensus about negative effects should be measured and reported, that will allow future studies to examine negative outcomes more precisely. Future research should also develop consensus among psychotherapy researchers that it is important to report negative outcomes in all trial reports, regardless of these are significant or not. Other research areas for future research include predictors of negative outcomes, the causes of negative outcomes and the pathways leading to them.

The results of this study should be considered in light of the limitations. The first limitation is that only a small fraction of trials on psychotherapy report deterioration rates. It is very well possible, therefore, that these studies are not a representative sample of all studies on psychotherapy for adult depression. However, a comparison between studies that did and those that did not report deterioration rates did not result in significant differences between the two samples of studies. A second limitation is the small number of trials and the considerable proportion of studies with risk of bias, making the results uncertain. Another limitation is that the studies used different ways to define deterioration, making it impossible to pool the prevalence rates across treatments and control groups.

Despite these limitations, this study showed that only a small number of studies have examined deterioration rates, that the studies that did report them found that deterioration is a relatively rare event, and that rates of deterioration may be lower in psychotherapy than in control conditions.

Funding sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflicts of interest

None.

Contributors

PC had the idea for this paper with input from all authors. PC also did the analyses and wrote the first draft. The searches were done by PC and EK, data extraction was done by PC, EK and MR. All authors have contributed to the next versions of the paper revising it critically for important intellectual content, and have agreed that this version is submitted for publication.

Acknowledgments

None.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jad.2018.05.050.

References

Barlow, D.H., 2010. Negative effects from psychological treatments: a perspective. Am. Psychol. 65, 13.
Barth, J., Munder, T., Gerger, H., Nuesch, E., Trelle, S., Znoj, H., Juni, P., Cuijpers, P., 2013. Comparative efficacy of seven psychotherapeutic interventions for patients with depression: a network meta-analysis. PLoS Med. 10, e1001454.
Boisvert, C.M., 2010. Negative treatment effects: is it time for a black box warning? Am. Psychol. 65, 680–681.
Buntrock, C., Ebert, D., Lehr, D., Riper, H., Smit, F., Cuijpers, P., Berking, M., 2015. Effectiveness of a web-based cognitive behavioural intervention for subthreshold depression: Pragmatic randomised controlled trial. Psychother. Psychosom. 84, 348–358.
Carlbring, P., Hagglund, M., Luthstrom, A., Dahlin, M., Kadowaki, A., Vernmark, K., Andersson, G., 2013. Internet-based behavioral activation and acceptance-based treatment for depression: a randomized controlled trial. J. Affect. Disord. 148, 331–337.
Choy, J.C., Lou, V.W., 2016. Effectiveness of the modified instrumental reminiscence intervention on psychological well-being among community-dwelling Chinese older adults: a randomized controlled trial. Am. J. Geriatr. Psychiat. 24, 60–69.
Cuijpers, P., 2017. Four decades of outcome research on psychotherapies for adult depression: an overview of a series of meta-analyses. Can. Psychol. 58, 7–19.
Cuijpers, P., Berking, M., Andersson, G., Quigley, L., Kleiboer, A., Dobson, K.S., 2013. A meta-analysis of cognitive-behavioural therapy for adult depression, alone and in comparison with other treatments. Can. J. Psychiatr. 58, 376–385.
Cuijpers, P., Cristea, I.A., Karyotaki, E., Reijnders, M., Huibers, M.J., 2016a. How effective are cognitive behavior therapies for major depression and anxiety disorders? A meta-analytic update of the evidence. World Psychiatr. 15, 245–258.
Cuijpers, P., Donker, T., Weissman, M.M., Ravitz, P., Cristea, I.A., 2016b. Interpersonal psychotherapy for mental health problems: a comprehensive meta-analysis. Am. J. Psychiatr. 173, 680–687.
Cuijpers, P., Driessen, E., Hollon, S.D., van Oppen, P., Barth, J., Andersson, G., 2012. The efficacy of non-directive supportive therapy for adult depression: a meta-analysis. Clin. Psychol. Rev. 32, 280–291.
Cuijpers, P., Karyotaki, E., Reijnders, M., Purgato, M., Barbui, C., 2018. Psychotherapies for depression in low- and middle-income countries: a meta-analysis. World Psychiatry. accepted for publication.
Cuijpers, P., van Straten, A., Warmerdam, L., Andersson, G., 2008b. Psychological treatment of depression: a meta-analytic database of randomized studies. BMC Psychiatr. 8, 36.
Dimidjian, S., Hollon, S.D., 2010. How would we know if psychotherapy were harmful? Am. Psychol. 65, 21–33.
Driessen, E., Hegelmaier, L.M., Abbass, A.A., Barber, J.P., Dekker, J.J.M., Van, H.L., Jansma, E.P., Cuijpers, P., 2015. The efficacy of short-term psychodynamic

psychotherapy for depression: a meta-analysis update. Clin. Psychol. Rev. 42, 1–15.
Duval, S., Tweedie, R., 2000. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometric 56, 455–463.
Ebert, D.D., Donkin, L., Andersson, G., Andrews, G., Berger, T., Carlbring, P., Cuijpers, P., 2016. Does Internet-based guided-self-help for depression cause harm? An individual participant data meta-analysis on deterioration rates and its moderators in randomized controlled trials. Psychol. Med. 46, 2679–2693.
Ekers, D., Webster, L., Van Straten, A., Cuijpers, P., Richards, D., Gilbody, S., 2014. Behavioural activation for depression; an update of meta-analysis of effectiveness and sub group analysis. PLoS One 9, e100100.
Foulkes, P., 2010. The therapist as a vital factor in side-effects of psychotherapy. Austr. N. Z. J. Psychiatr. 44 189–189.
Gitlin, L.N., Harris, L.F., McCoy, M.C., Chernett, N.L., Pizzi, L.T., Jutkowitz, E., Hess, E., Hauck, W.W., 2013. A home-based intervention to reduce depressive symptoms and improve quality of life in older African Americans: a randomized trial. Ann. Intern. Med 159, 243–252.
Hadley, S.W., Strupp, H.H., 1976. Contemporary views of negative effects in psychotherapy: an integrated account. Arch. Gen. Psychiatr. 33, 1291–1302.
Hallford, D.J., Mellor, D., 2016. Autobiographical memory-based intervention for depressive symptoms in young adults: a randomized controlled trial of cognitive-reminiscence therapy. Psychoth. Psychosom. 85, 246–249.
Hautzinger, M., Welz, S., 2004. Cognitive behavioral therapy for depressed elderly outpatients - a controlled randomized trial. Verhaltensth. Psychosoz. Prax. 36, 789–798.
Higgins, J.P.T., Altman, D.G., Gøtzsche, P.C., Jüni, P., Moher, D., Oxman, A.D., Savović, J., Schulz, K.F., Weeks, L., Sterne, J.A.C., 2011. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. B.M.J 343, d5928.
Higgins, J.P.T., Thompson, S.G., Deeks, J.J., Altman, D.G., 2003. Measuring inconsistency in meta-analyses. B.M.J. 327, 557–560.
Jacobson, N.S., Truax, P., 1991. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. J. Consult. Clin. Psychol. 59, 12–19.
Ioannidis, J.P.A., Patsopoulos, N.A., Evangelou, E., 2007. Uncertainty in heterogeneity estimates in meta-analyses. B.M.J 335, 914–916.
Kelly, J.A., Murphy, D.A., Bahr, G.R., Kalichman, S.C., Morgan, M.G., Stevenson, L.Y., Koob, J.J., Brasfield, T.L., Bernstein, B.M., 1993. Outcome of cognitive-behavioral and support group brief therapies for depressed, HIV-infected persons. Am. J. Psychiatr. 150, 1679–1686.
Lambert, M., 2007. Presidential address: what we have learned from a decade of research aimed at improving psychotherapy outcome in routine care. Psychoth. Res. 17, 1–14.
Lappalainen, P., Langrial, S., Oinas-Kukkonen, H., Tolvanen, A., Lappalainen, R., 2015. Web-based acceptance and commitment therapy for depressive symptoms with minimal support: a randomized controlled trial. Beh. Mod. 39, 805–834.
Lilienfeld, S.O., 2007. Psychological treatments that cause harm. Persp. Psychol. Sci. 2, 53–70.
Linden, M, Schermuly-Haupt, M.L., 2014. Definition, assessment and rate of psychotherapy side effects. World Psychiatry 13, 306–309.
Malouff, J.M., Thorsteinsson, E.B., Schutte, N.S., 2007. The efficacy of problem solving therapy in reducing mental and physical health problems: a meta-analysis. Clin. Psychol. Rev. 27 (1), 46–57.
Mohr, D.C., 1995. Negative outcome in psychotherapy: a critical review. Clin. Psychol. 2, 1–27.
Mulcahy, R., Reay, R.E., Wilkinson, R.B., Owen, C., 2010. A randomised control trial for the effectiveness of group Interpersonal Psychotherapy for postnatal depression. Arch. Womens Ment. Health. 13, 125–139.
Nobis, S., Lehr, D., Ebert, D.D., Baumeister, H., Snoek, F., Riper, H., Berking, M., 2015. Efficacy of a web-based intervention with mobile phone support in treating depressive symptoms in adults with type 1 and type 2 diabetes: a randomized controlled trial. Diabetes Care 38, 776–783.
Orsini, N., Bottai, M., Higgins, J., Buchan, I., 2006. Heterogi: stata module to quantify heterogeneity in a meta-analysis. Stat. Softw. Compon.
Peterson, A.L., Roache, J.D., Raj, J., Young-McCaughan, S., Consortium, S.S., 2013. The need for expanded monitoring of adverse events in behavioral health clinical trials. Contemp. Clin. Trials 34, 152–154.
Pugh, N.E., Hadjistavropoulos, H.D., Hampton, A.J.D., Bowen, A., Williams, J., 2015. Client experiences of guided internet cognitive behavior therapy for postpartum depression: a qualitative study. Arch. Women's Ment. Health. 18, 209–219.
Ransom, D., Heckman, T.G., Anderson, T., Garske, J., Holroyd, K., Basta, T., 2008. Telephone-delivered, interpersonal psychotherapy for HIV-infected rural persons with depression: a pilot trial. Psychiatr. Serv. 59, 871–877.
Rozental, A., Andersson, G., Boettcher, J., Ebert, D.D., Cuijpers, P., Knaevelsrud, C., Ljótsson, B., Kaldo, V., Titov, N., Carlbring, P., 2014. Consensus statement on defining and measuring negative effects of internet interventions. Internet Interv. 1, 12–19.
Scott, M.J., Stradling, S.G., 1990. Group cognitive therapy for depression produces clinically significant reliable change in community-based settings. Beh. Cogn. Psychoth. 18, 1–19.
Segre, L.S., Brock, R.L., O'Hara, M.W., 2015. Depression treatment for impoverished mothers by point-of-care providers: a randomized controlled trial. J. Cons. Clin. Psychol. 83, 314–324.
Vittengl, J.R., Jarrett, R.B., Weitz, E., Hollon, S.D., Twisk, J., Cristea, I., …, Cuijpers, P., 2016. Divergent outcomes in cognitive-behavioral therapy and pharmacotherapy for adult depression. Am. J. Psychiatr. 173, 481–490.
Watt, L.M., Cappeliez, P., 2000. Integrative and instrumental reminiscence therapies for depression in older adults: intervention strategies and treatment effectiveness. Aging Ment. Health. 4, 166–177.
Wickberg, B., Hwang, C.P., 1996. Counselling of postnatal depression: a controlled study on a population based Swedish sample. J. Affect. Disord. 39, 209–216.
Wollersheim, J.P., Wilson, G.L., 1991. Group treatment of unipolar depression: a comparison of coping, supportive, bibliotherapy, and delayed treatment groups. Prof. Psychol. 22, 496.
Wuthrich, V.M., Rapee, R.M., 2013. Randomised controlled trial of group cognitive behavioural therapy for comorbid anxiety and depression in older adults. Beh. Res. Ther. 51, 779–786.